On the Optimality of the Regular Simplex Code

By H. J. LANDAU and DAVID SLEPIAN

(Manuscript received May 25, 1966)

We prove here the long conjectured fact that the regular simplex is the code of minimal error probability for transmission over the infinite-band Gaussian channel. The code is actually optimal for a rather wide class of assumed channel noises. We also establish the optimality of several other codes for the band-limited Gaussian channel.

I. INTRODUCTION

Since its introduction by Shannon[1] and Kotel'nikov[2] nearly 20 years ago, the geometric representation of signals has played an important role in communication theory.* By this scheme, a variety of physically different time-continuous communication systems can all be reduced to the same geometric model. The problem of finding optimal signals for transmission then becomes a geometric one. Tbis paper solves one such problem.

In the model in question, signals to be transmitted are represented as points, or vectors from the origin, in a suitable finite dimensional Euclidean signal space $\mathcal{E}_n$. The energy of any signal in $\mathcal{E}_n$ is proportional to the length of its representative vector; the bandwidtb of tbe communication system is proportional to the dimension n of the signal space. Received signals are also represented by vectors in $\mathcal{E}_n$ and tbe difference $\mathbf{Z} = \mathbf{Y} - \mathbf{X}$ between a transmitted signal $\mathbf{X}$ and tbe corresponding received signal $\mathbf{Y}$ is a vector random variable representative of the noise encountered during transmission. In a model commonly considered, the probability density of $\mathbf{Z}$ depends only on its magnitude, i.e.,

$$p(z_1, z_2, \cdots, z_n) = f(|\mathbf{Z}|), \tag{1}$$

* A detailed description of this viewpoint along with some references to the intervening literature can be found in Chapters 4 and 5 of the recent book[3] by Wozencraft and Jacobs.

and $f(\cdot)$ is an integrable nonnegative monotone decreasing function of its argument. We shall consider only this case in all that follows.

Suppose the transmitter has a list of M signals, $\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_M$ from which it selects the successive signals to be transmitted. We suppose these choices are made independently with equal probabilities and that the code, or list of possible sent signals, is known to the receiver. The receiver partitions $\mathcal{E}_n$ into M disjoint regions $\mathcal{R}_1, \mathcal{R}_2, \cdots, \mathcal{R}_M$ called decision regions. When the received signal lies in $\mathcal{R}_i$, the receiver asserts that $\mathbf{X}_i$ was transmitted. With this scheme, the probability of correct decoding is

$$Q = \frac{1}{M} \sum_{i=1}^{M} \int_{\mathcal{R}_i} f(|\mathbf{X} - \mathbf{X}_i|)\, dx_1 dx_2 \cdots dx_n, \tag{2}$$

where $\mathbf{X} = (x_1, x_2, \cdots, x_n)$ is a generic point in $\mathcal{E}_n$.

With M and n given, how large can Q be made by proper choice of the code and decision regions? For a given code it is well known (see Ref. 3, Section 4.2, for example) that Q is maximized by choosing

$$\mathcal{R}_i = \{\mathbf{X} \mid |\mathbf{X} - \mathbf{X}_i| < |\mathbf{X} - \mathbf{X}_j|, \quad j \neq i\}, \tag{3}$$

$i = 1, 2, \cdots, M$. That is, the ith decision region consists of all points of $\mathcal{E}_n$ closer to $\mathbf{X}_i$ than to any other code word. Decision regions determined by (3) are known as maximum-likelihood regions.

The maximization of Q over the code is more complicated. To obtain a meaningful problem it is necessary to put some restriction on the length of the code vectors, for without this, Q can be made arbitrarily close to unity by choosing large enough vectors in distinct directions. Several different energy restrictions have been studied in the literature (see Ref. 4). Although optimal codes under these restrictions have not been found in general for fixed M and n, much detail is known in the Gaussian case

$$f(x) = \frac{\exp(-x^2/2\sigma^2)}{\sqrt{2\pi\sigma^2}}, \tag{4}$$

about the asymptotic form of Q for such optimal codes, as $n \to \infty$ with $(1/n) \log M \to R$. These results are usually described in the channel capacity and reliability formulae terms of information theory.[3,4,5]

In this paper we restrict our attention to the case in which all code vectors are the same length. For convenience, we take

$$|\mathbf{X}_i| = 1, \qquad i = 1, 2, \cdots, M. \tag{5}$$

Such codes are called "equal energy codes".[6] The code optimization problem can then be stated as follows. Find M points $\hat{\mathbf{X}}_1, \hat{\mathbf{X}}_2, \cdots, \hat{\mathbf{X}}_M$

on the unit sphere* in $\mathcal{E}_n$ such that Q as given by (2) and (3) attains its maximum value.

To our knowledge, the first investigation of particular codes from this geometric point of view was carried out in 1948 by L. A. MacColl[7] who investigated codes corresponding to the vertices of the three regular polytopes[8] in n-space. These are the regular simplex for which $M = n + 1$, the hypercube for which $M = 2^n$ and the cross-polytope or biorthogonal code for which $M = 2n$. MacColl wrote explicit expressions for Q for these codes and evaluated them numerically for the Gaussian case (4) for a variety of values of n and σ. Gilbert[9] continued this work and made comparisons with a variety of other point configurations. Balakrishnan[10] established a new expression for Q in the Gaussian case, which permitted him to show that the regular simplex code is locally optimal (yields a larger Q than nearby equal energy codes with $M = n + 1$). Later [11] he showed that as $\sigma \to \infty$ and as $\sigma \to 0$ the optimal code of $n + 1$ points approached the regular simplex. Weber[12] used Balakrishnan's form for Q to show that for $n = 2$ the (globally) optimal code of M points, $M = 3, 4, \cdots$, is the regular M-gon. For $n = 2, 3, \cdots$, he also showed the biorthogonal code to be a local optimum among equal energy codes with $M = 2n$, and described a family of locally optimal codes for $M = n + 1, n + 2, \cdots, 2n$.

In this paper, we at last lay to rest the longstanding conjecture that the regular simplex is optimal for $M = n + 1$ in the Gaussian case.† Specifically, we show that Q as given by (2)–(3) is greater for the regular simplex than for any other equal energy code of $M = n + 1$ points in $\mathcal{E}_n$, $n = 3, 4, 5, \cdots$. This result is true for any monotone decreasing f. The method of proof is based on a generalization to higher dimensions of a theorem of Fejes-Tóth[13] concerning expressions related to the form (2) in 3 dimensions.‡ Our methods also establish that the optimal equal energy codes with parameters $M = 6$, $n = 3$, and $M = 12$, $n = 3$ are, respectively, the biorthogonal code and the code consisting of the midpoints of the faces of the regular dodecahedron. We conclude with some comments about the biorthogonal code and about the reliability of the infinite-band Gaussian channel.

II. AN INEQUALITY FOR Q

For an equal energy code, the maximum-likelihood region $\mathcal{R}_i$ given by (3) can be determined as follows. Let $\mathcal{H}_{ij}$ denote the hyperplane that

* We shall hereafter use the caret ^ to denote unit vectors.
† It is incorrectly stated in Ref. 3, pp. 260, 364 that this result has been previously shown in the literature.
‡ We are indebted to E. N. Gilbert for calling Fejes-Tóth's work to our attention.

bisects perpendicularly the line segment joining $\hat{\mathbf{X}}_i$ to $\hat{\mathbf{X}}_j$. This plane passes through the origin and divides $\mathcal{E}_n$ into two half-spaces. We denote by $\mathcal{U}_{ij}$ the half-space containing $\hat{\mathbf{X}}_i$. It consists of all points of $\mathcal{E}_n$ closer to $\hat{\mathbf{X}}_i$ than to $\hat{\mathbf{X}}_j$. The region $\mathcal{R}_i$ is the intersection of M-1 such half-spaces,

$$\mathcal{R}_i = \bigcap_{\substack{j=1\\ j\neq i}}^{M} \mathcal{U}_{ij}.$$

It is, therefore, a convex region bounded by a certain number of hyperplane faces that pass through the origin — a kind of flat-sided cone with vertex at the origin. We note that the various maximum-likelihood regions, $\mathcal{R}_1, \mathcal{R}_2, \cdots, \mathcal{R}_M$, are disjoint and that together with their boundaries they exhaust $\mathcal{E}_n$.

Let us now call any convex region of $\mathcal{E}_n$ bounded by $k \geqq n$ hyperplanes through the origin a "flat-sided cone". We shall establish an upper bound for Q as given by (2) when the M decision regions $\mathcal{R}_i$ are any set of disjoint flat-sided cones (not necessarily maximum-likelihood regions of any code) that together with their boundaries exhaust $\mathcal{E}_n$. For our purposes it suffices to consider only the case in which $\hat{\mathbf{X}}_i$ lies in the interior of $\mathcal{R}_i$, $i = 1, 2, \cdots, M$.

We denote by S the surface of the unit sphere in $\mathcal{E}_n$ with center at the origin. We denote by R_i the intersection of $\mathcal{R}_i$ with S. The regions R_i are "spherical polygons" that reticulate S into a map or net. We shall evaluate Q by first integrating over this net on S and by then performing a radial integration.

Let $\mathbf{X}$ be a generic point in $\mathcal{E}_n$ distant r from the origin (see Fig. 1)

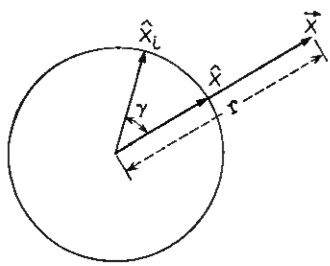

Fig. 1 — Reduction to unit sphere.

and let $\hat{\mathbf{X}}$ be a unit vector in the direction of $\mathbf{X}$, i.e., the terminus of $\hat{\mathbf{X}}$ is the radial projection of the generic point onto S. Then

$$|\hat{\mathbf{X}}_i - \mathbf{X}|^2 = 1 + r^2 - 2r\cos\gamma$$

and

$$|\hat{\mathbf{X}}_i - \hat{\mathbf{X}}|^2 = 2 - 2\cos\gamma$$

so that

$$|\hat{\mathbf{X}}_i - \mathbf{X}|^2 = (1-r)^2 + r\,|\hat{\mathbf{X}}_i - \hat{\mathbf{X}}|^2.$$

We can thus write $f(\,|\hat{\mathbf{X}}_i - \mathbf{X}|\,) = g_r(\,|\hat{\mathbf{X}}_i - \hat{\mathbf{X}}|\,)$ where for each fixed r the function $g_r(\cdot)$ is nonnegative and is monotone decreasing in its argument. The expression (2) in these terms becomes

$$Q = \int_0^\infty dr\, r^{n-1} U(r) \tag{6}$$

$$U(r) = \frac{1}{M} \sum_{i=1}^{M} \int_{R_i} g_r(|\hat{\mathbf{X}}_i - \hat{\mathbf{X}}|)ds \tag{7}$$

where ds is the differential surface- or $(n-1)$-content of S at the point $\hat{\mathbf{X}}$. Note that $U \geqq 0$. We proceed to find an upper bound for U. By (6) this will provide the desired bound for Q.

Let the terminus of the unit vector $\hat{\mathbf{Y}}$ determine a point P on S (see Fig. 2). The set of all points $\hat{\mathbf{X}}$ on S such $\hat{\mathbf{X}}\cdot\hat{\mathbf{Y}} \geqq \cos\varphi \geqq 0$ will be called "the spherical cap of S of angle φ about P". Now let $\mathcal{H}$ be a hyperplane through the origin but not containing P that intersects this spherical cap. That is $0 < \hat{\mathbf{n}}\cdot\hat{\mathbf{Y}} < \sin\varphi$ where $\hat{\mathbf{n}}$ is the unit normal to $\mathcal{H}$ directed positively toward the side on which P lies. $\mathcal{H}$ divides the spherical cap into two parts. We denote by W the part of the cap not containing P, and we denote by w the content of W.

In what follows, the function

$$h_r(w) = \int_W g_r(|\hat{\mathbf{Y}} - \hat{\mathbf{X}}|)ds \tag{8}$$

will be of great importance to us. The notation suppresses the dependence of h on φ, the angle of the spherical cap, and points out that with the

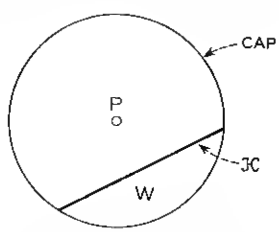

Fig. 2 — Cap cut by hyperplane.

geometry as described this integral is a function only of

$$w = \int_W ds \tag{9}$$

the content of W. We shall suppose φ fixed in all that follows.

Two special properties of $h_r(w)$ are of particular concern. First, as shown in Appendix B, this function is increasing and convex. That is, if $w_2 > w_1$, then $h_r'(w_2) > h_r'(w_1)$ where $h_r'(w) = dh_r/dw$. This, of course, implies that

$$\sum p_i h_r(w_i) \geqq h_r(\sum p_i w_i), \tag{10}$$

where the p_i are nonnegative weight factors summing to unity. Equality holds only when the w_i are all equal.

The second property of $h_r(w)$ is somewhat more complicated to state, though in three dimensions it is intuitively obvious. Again let $\mathcal{H}$ be a hyperplane through the origin but not through P that cuts off a piece W of the spherical cap about P. Let $\mathcal{H}_1, \mathcal{H}_2, \cdots, \mathcal{H}_j$ be hyperplanes that each contain the origin and P. We denote by V the portion of W lying on the positive side of $\mathcal{H}_i$, $i = 1, 2, \cdots, j$, and we denote the content of V by v. It is established in Appendix C that

$$\int_V g_r(|\hat{\mathbf{Y}} - \hat{\mathbf{X}}|)ds \geqq h_r(v), \tag{11}$$

where as before $\hat{\mathbf{Y}}$ is the vector from the origin to P and $\hat{\mathbf{X}}$ is a generic point of V. Equality holds only if $\mathcal{H}$ is the sole hyperplane boundary of V (i.e., if none of $\mathcal{H}_1, \mathcal{H}_2, \cdots, \mathcal{H}_j$ form a part of the boundary of V).

With these two properties of $h_r(w)$ we can now establish the desired inequality for $U(r)$. We first "triangulate" each of the polygonal regions R_i into "spherical pyramids" R_{ij} having boundaries of R_i as bases and $\hat{\mathbf{X}}_i$ as a vertex. More accurately described, the regions R_{ij} are found as follows. The flat-sided cone $\mathcal{R}_i$ is bounded by pieces of k_i (say) hyperplanes $\mathcal{H}_1^{(i)}, \cdots, \mathcal{H}_{k_i}^{(i)}$ through the origin. We denote by $\mathcal{B}_{ij}$ the portion of $\mathcal{H}_j^{(i)}$ that bounds $\mathcal{R}_i$. Now $\mathcal{B}_{ij}$ is itself bounded by a certain number l_{ij} of $(n-2)$-flats through the origin. Through each of these $(n-2)$-flats we pass a hyperplane $\mathcal{H}_k^{(ij)}$, $k = 1, 2, \cdots, l_{ij}$, that contains $\hat{\mathbf{X}}_i$. These hyperplanes, along with $\mathcal{B}_{ij}$, determine a new flat-sided cone $\mathcal{R}_{ij}$ having $\mathcal{B}_{ij}$ as one face and the line containing $\hat{\mathbf{X}}_i$ as a one-dimensional boundary. The interiors of the k_i flat-sided cones $\mathcal{R}_{i1}, \mathcal{R}_{i2}, \cdots, \mathcal{R}_{ik_i}$ are disjoint. Together with their boundaries they exhaust $\mathcal{R}_i$. The line through $\hat{\mathbf{X}}_i$ is common to the boundaries of all k_i of these flat-sided cones. The spherical pyramid R_{ij} is the intersection of $\mathcal{R}_{ij}$ with S.

We denote by C_i the spherical cap of S of angle φ about $\hat{\mathbf{X}}_i$ (see Fig. 3).

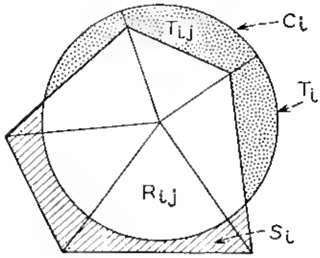

Fig. 3 — Cap and triangulated decision region.

Let T_i be the portion of C_i exterior to R_i and let S_i be the portion of R_i exterior to C_i. A typical term of the sum (7) can then be written

$$\int_{R_i} g_r(|\hat{\mathbf{X}}_i - \hat{\mathbf{X}}|)ds = \left(\int_{C_i} + \int_{S_i} - \int_{T_i}\right) g_r(|\hat{\mathbf{X}}_i - \hat{\mathbf{X}}|)ds.$$

Now T_i can be broken up into pieces T_{ij} corresponding to the spherical pyramids R_{ij}. To accomplish this, we extend the sides of the pyramid beyond its base. Thus, if $\mathcal{R}_{ij}$ is on the positive side of $\mathcal{H}_j^{(i)}$ and $\mathcal{H}_k^{(ij)}$, $k = 1, 2, \cdots, l_{ij}$, then T_{ij} is the part of the spherical cap on the negative side of $\mathcal{H}_j^{(i)}$ and the positive side of $\mathcal{H}_k^{(ij)}$, $k = 1, 2, \cdots, l_{ij}$. We now have

$$\int_{R_i} g_r ds = \int_{C_i} g_r ds + \int_{S_i} g_r ds - \sum_{j=1}^{k_i} \int_{T_{ij}} g_r ds. \tag{12}$$

Some of the regions S_i, T_{ij} can, of course, be void.

We now sum (12) over the M regions R_i. We write

$$k = \frac{1}{2} \sum_{i=1}^{M} k_i$$

for the total number of $(n - 2)$-boundaries in the net on S. (Each boundary of R_i is shared with one other spherical polygon.) There results

$$\begin{aligned} MU(r) &= \sum_{i=1}^{M} \int_{R_i} g_r(|\hat{\mathbf{X}}_i - \hat{\mathbf{X}}|)ds = M \int_{C_1} g_r(|\hat{\mathbf{X}}_1 - \hat{\mathbf{X}}|)ds \\ &\quad + \sum_{i=1}^{M} \int_{S_i} g_r(|\hat{\mathbf{X}}_i - \hat{\mathbf{X}}|)ds - \sum_{i=1}^{M} \sum_{j=1}^{k_i} \int_{T_{ij}} g_r(|\hat{\mathbf{X}}_i - \hat{\mathbf{X}})ds. \end{aligned} \tag{13}$$

We next use (11) for the regions T_{ij}.

$$MU(r) \leqq M \int_{C_1} g_r(|\hat{\mathbf{X}}_1 - \hat{\mathbf{X}}|)ds + \sum_{i=1}^{M} \int_{S_i} g_r ds - \sum_{i=1}^{M} \sum_{j=1}^{k_i} h_r(t_{ij})$$

where t_{ij} is the content of T_{ij}. The convexity (10) of h now gives

$$MU(r) \leqq M \int_{C_1} g_r(|\hat{\mathbf{X}}_1 - \hat{\mathbf{X}}|)ds + \sum_{i=1}^{M} \int_{S_i} g_r ds - 2kh_r \left(\frac{1}{2k} \sum_{i=1}^{M} \sum_{j=1}^{k_i} t_{ij} \right). \tag{14}$$

Now denote by s, c, and s_i, respectively, the content of S, C_i, and S_i. A division of S analogous to (13) (set $g = 1$ there) gives

$$\begin{aligned} s &= Mc + \sum_{1}^{M} s_i - \sum_{i=1}^{M} \sum_{j=1}^{k_i} t_{ij} \\ &= Mc + s' - \sum_{i=1}^{M} \sum_{j=1}^{k_i} t_{ij} \end{aligned} \tag{15}$$

where we write $s' = \sum s_i$ for the sum of the contents of all of the pieces of the polygons R_i that fall outside their respective spherical caps. We then have from (14) and (15)

$$\begin{aligned} MU(r) &\leqq M \int_{C_1} g_r ds + \sum_{i=1}^{M} \int_{S_i} g_r ds - 2kh_r \left(\frac{Mc - s + s'}{2k} \right) \\ &= M \int_{C_1} g_r ds - 2kh_r \left(\frac{Mc - s}{2k} \right) \\ &\quad + \sum_{i=1}^{M} \int_{S_i} g_r ds - 2k \int_{K} g_r(|\hat{\mathbf{Y}} - \hat{\mathbf{X}}|)ds \end{aligned} \tag{16}$$

where K is the region (see Fig. 4) of the spherical cap about P that lies between hyperplanes through the origin that cut from the cap regions of content $(Mc - s + s')/2k$ and $(Mc - s)/2k$. This latter quantity will henceforth be assumed to be nonnegative. The normals to the two hyperplanes and the vector $\hat{\mathbf{Y}}$ from the origin to P are chosen coplanar.

Note now that the sum of tbe last two terms in (16) cannot be positive, for we have

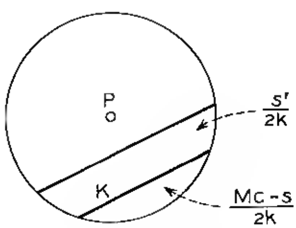

Fig. 4 — The region K.

$$\sum_{i=1}^{M} \int_{S_i} g_r(|\hat{\mathbf{X}}_i - \hat{\mathbf{X}}|)ds \leqq g_r(d) \sum_{i=1}^{M} \int_{S_i} dr = g_r(d)s' \tag{17}$$

and

$$2k \int_K g_r(|\hat{\mathbf{Y}} - \hat{\mathbf{X}}|)ds \geqq g_r(d)2k \int_K ds = g_r(d)s' \tag{18}$$

where $d = \sqrt{2 - 2\cos\varphi}$ is the distance from the center of the cap to the edge. Here we have used the fact that g_r is monotone decreasing. Equality holds in (17) and (18) only when $s' = 0$.

In view of the above, the inequality in (16) can be continued by omitting the last two terms there and we then obtain our desired inequality

$$MU(r) \leqq M \int_C g_r(|\hat{\mathbf{Y}} - \hat{\mathbf{X}}|)ds - 2kh_r\left(\frac{Mc - s}{2k}\right) \tag{19}$$

where C is the spherical cap of angle φ about the terminus of $\hat{\mathbf{Y}}$ and we require $Mc - s \geqq 0$. Retracing all the inequalities used to derive (19), we see that the equality sign holds there if and only if $s' = 0$, all the regions T_{ij} have equal content and each T_{ij} is a region cut off from the spherical cap by a single hyperplane.

In closing this section we note one further fact. From the convexity of $h_r(x)$ it follows that $xh_r(\alpha/x)$ is monotone decreasing in x. For given M and c, then, the right side of (19) is monotone increasing in k.

III. OPTIMALITY OF THE REGULAR SIMPLEX AND CERTAIN OTHER CODES

Let a code of M unit vectors in $\mathcal{E}_n$ have maximum-likelihood regions $\mathcal{R}_i$ that reticulate the surface of the unit sphere into a net having k $(n - 2)$-dimensional boundaries. We designate such a code by the symbol $\{n,M,k\}$. For certain values of the parameters n, M and k, there may exist codes for which a spherical cap angle φ can be found such that the conditions for equality hold in (19). We call such a code a *symmetric* $\{n,M,k\}$. By choosing C so that (19) is an equality for such a symmetric code, we see that the probability Q of no error for a symmetric $\{n,M,k\}$ is greater than the no-error probability of any nonsymmetric $\{n,M,k\}$. Indeed, the concluding remark of Section II shows that the no-error probability of a symmetric $\{n,M,k\}$ is greater than the no-error probability of every $\{n,M,k'\}$ if $k' < k$.

The regular simplex code consists of $M = n + 1$ unit vectors $\hat{\mathbf{X}}_1, \hat{\mathbf{X}}_2, \cdots, \hat{\mathbf{X}}_{n+1}$ in $\mathcal{E}_n$ with $\hat{\mathbf{X}}_i \cdot \hat{\mathbf{X}}_j = -(1/n)$, $i \neq j$. The maximum-likelihood region $\mathcal{R}_i$ containing $\hat{\mathbf{X}}_i$ is bounded by n hyperplanes. It is

readily verified that the regular simplex is a symmetric $\{n, n+1, n(n+1)/2\}$. Now no code of $n+1$ unit vectors in $\mathcal{E}_n$ can have more than $k = n(n+1)/2$ $(n\text{-}2)$-boundaries in its maximum-likelihood net, for, by the construction given in the first paragraph of Section II, each maximum-likelihood region can be bounded by at most n hyperplanes. The regular simplex code then must have a Q strictly greater than any other equal energy code of $n+1$ vectors in $\mathcal{E}_n$ except possibly another distinct symmetric $\{n, n+1, n(n+1)/2\}$, should such exist. But this latter eventuality cannot happen. That a symmetric $\{n, n+1, n(n+1)/2\}$ must be the regular simplex can be seen as follows. Since no $\mathcal{R}_i$ for a code of $M = n+1$ points can have more than n hyperplane boundaries, then to have $k = n(n+1)/2$, every $\mathcal{R}_i$ must have exactly n hyperplane boundaries. The n hyperplanes $\mathcal{H}_1^{(1)}, \mathcal{H}_2^{(1)}, \cdots, \mathcal{H}_n^{(1)}$ that bound $\mathcal{R}_1$ bisect, respectively, the line segments from $\hat{\mathbf{X}}_1$ to $\hat{\mathbf{X}}_2$, from $\hat{\mathbf{X}}_1$ to $\hat{\mathbf{X}}_3, \cdots$, from $\hat{\mathbf{X}}_1$ to $\hat{\mathbf{X}}_{n+1}$. Since the code is assumed symmetric, these hyperplanes must be equidistant from $\hat{\mathbf{X}}_1$. Thus, all the other code points are equidistant from $\hat{\mathbf{X}}_1$. But a similar argument holds for each of the other regions $\mathcal{R}_2, \mathcal{R}_3, \cdots, \mathcal{R}_{n+1}$ and so all distances between pairs of code points are equal. But this property suffices to define the regular simplex.

The optimality of two other codes in $n = 3$ dimensions can readily be established by using (19). We note first that in 3 dimensions the conditions for equality to hold in (19) are such that the maximum-likelihood net on the sphere must be composed of congruent regular spherical polygons. A symmetric $\{3,M,k\}$ then must be the radial projection onto the unit sphere of a regular three-dimensional polyhedron. The code points are the centers of the faces of the polyhedron.

Consider now the code formed by the midpoints of the faces of a cube of edge length 2. This is the three-dimensional biorthogonal code. The maximum-likelihood net is given by the radial projection of the cube edges onto the inscribed unit sphere. The code is a symmetric $\{3,6,12\}$. There is no regular polyhedron with 6 faces other than the cube, so that we will have shown the three-dimensional biorthogonal code to be optimal if we establish that every $\{3,6,k\}$ must have $k \leqq 12$. To see this latter fact, note that for three-dimensional codes, at least three edges of the maximum-likelihood net must meet at each vertex of the net (since each R_i is convex). Thus $3v \leqq 2k$ where v and k are, respectively, the total number of vertices and edges for the net. Euler's formula (Ref. 8, p. 9) $v - k + M = 2$ holds for the net, and so

$$k \leqq 3(M-2). \tag{20}$$

For the case at hand $M = 6$, and (20) gives $k \leqq 12$, so that the proof is completed.

Analogous reasoning shows that the centers of the faces of the dodecahedron give the best code with $M = 12$ points. The code is a symmetric $\{3,12,30\}$.

The regular octahedron in $\mathcal{E}_3$ gives rise to a symmetric $\{3,8,12\}$ whose code points are the vertices of a cube. This is not the optimal configuration of 8 points in $\mathcal{E}_3$. By rotating one face of the cube 45 degrees about an axis perpendicular to the face and through its center, one obtains a $\{3,8,16\}$. By translating this face slightly toward the opposite face of the cube, and by slightly expanding both faces, one obtains a $\{3,8,16\}$ with minimum distance between code points strictly larger than the minimum for the cubic arrangement of points. There are then noise functions $f(\mid z \mid)$ of (1) for which this new code has a larger Q than the cube-code.

It is not known whether the symmetric $\{3,20,30\}$ obtained from the regular icosahedron is an optimal code of 20 points.

IV. THE BIORTHOGONAL CODE

The biorthogonal code is a symmetric $\{2n,n,n(2n - 2)\}$. The $2n$ code points can be taken as the points on the coordinate axes unit distance from the origin. Alternatively, the code points can be described as the centers of the $(n - 1)$-dimensional bounding cells of the unit n-cube. The radial projection of the cube onto the unit sphere with center at the center of the hypercube gives the maximum-likelihood net of the code.

We have seen that for $n = 3$ the biorthogonal code is optimal among codes of $M = 2n = 6$ points. It is natural to suspect that for all n the biorthogonal code is optimal among codes of $2n$ points in $\mathcal{E}_n$. However, the methods used in this paper, based as they are on (19), will not suffice to settle this question, for, as will be shown below, when $n \geqq 4$, there exist $\{2n,n,n(2n - 1)\}$ codes; i.e., codes with a larger k value than the biorthogonal code.

It might be thought that this encumbering dependence of (19) on k could be avoided — that an inequality for Q independent of k could be found which is attainable for optimal codes. The example already treated of the octahedron shows, however, that this dependence on k is essential.

To construct a $\{2n,n,n(2n - 1)\}$ for $n \geqq 5$, choose $2n$ distinct real numbers $\nu_1, \nu_2, \cdots, \nu_{2n}$. The vectors of the code are given by

$$\hat{\mathbf{X}}_i = (\alpha_i, \alpha_i\nu_i, \alpha_i\nu_i^2, \cdots, \alpha_i\nu_i^{n-1}), \tag{21}$$

where

$$\alpha_i = [1 + \nu_i^2 + \nu_i^4 + \cdots + \nu_i^{2n-2}]^{-\frac{1}{2}} \qquad i = 1, 2, \cdots, 2n$$

has been chosen so that $\hat{\mathbf{X}}_i$ is a unit vector. The code is closely related to the cyclic polytope described by Gale[14].

An important property of this code can be derived[14] by considering the polynomials

$$F_{ij}(\lambda) \equiv (\lambda - \nu_i)^2(\lambda - \nu_j)^2 = \sum_{p=0}^{4} A_{ij}^{(p)}\lambda^p \qquad i,j = 1, 2, \cdots, 2n \tag{22}$$

which are nonnegative. We define the $(2n)^2$ n-vectors

$$\mathbf{B}_{ij} \equiv (A_{ij}^{(0)}, A_{ij}^{(1)}, A_{ij}^{(2)}, A_{ij}^{(3)}, A_{ij}^{(4)}, 0, \cdots, 0).$$

We then have

$$\begin{aligned} \mathbf{B}_{ij}\cdot\hat{\mathbf{X}}_l &= \alpha_l \sum_{p=0}^{4} A_{ij}^{(p)}\nu_l^p = F_{ij}(\nu_l) \\ &= \begin{cases} 0, & l = i \\ 0, & l = j \\ a_{ijl} > 0, & l \neq i, \quad l \neq j \end{cases} \end{aligned} \tag{23}$$

where the positivity of the a_{ijl} follows from the factored form (22) of $F_{ij}(\lambda)$.

To show that the points (21) determine a $\{2n,n,n(2n - 1)\}$ we note first that they span $\mathcal{E}_n$. Indeed every choice of n vectors $\hat{\mathbf{X}}_i$ from (21) yields an independent set, as can be seen by forming the determinant whose rows are the components of the vectors. These determinants are proportional to Vandermonde determinants and do not vanish. To show that $k = n(2n - 1)$ for the code, consider the maximum likelihood region $\mathcal{R}_i$ containing $\hat{\mathbf{X}}_i$. By the construction described in the first paragraph of Section II, $\mathcal{R}_i$ is the intersection of the half-spaces

$$\mathcal{H}_j^{(i)}(\mathbf{X}) = (\hat{\mathbf{X}}_i - \hat{\mathbf{X}}_j)\cdot\hat{\mathbf{X}} \geqq 0 \qquad j = 1, 2, \cdots, 2n; \qquad j \neq i. \tag{24}$$

We assert that each of the $2n - 1$ hyperplanes $\mathcal{H}_j^{(i)}$, $j = 1,2, \cdots, 2n$ with $j \neq i$, is indeed an $(n - 1)$-dimensional boundary of $\mathcal{R}_i$. It will then follow that $k = \frac{1}{2}2n(2n - 1)$ since there are $2n$ maximum likelihood regions. That $\mathcal{H}_j^{(i)}$ is an $(n - 1)$-boundary of $\mathcal{R}_i$ results from the fact that there exists a point $\mathbf{X}_0$ contained in $\mathcal{R}_i$ that lies in $\mathcal{H}_j^{(i)}$ but not in $\mathcal{H}_k^{(i)}$, $k = 1, 2, \cdots, 2n$ with $k \neq i$ and $k \neq j$. From (23) we can choose $\mathbf{X}_0 = \mathbf{B}_{ij}$ since

$$\mathcal{H}_j^{(i)}(\mathbf{B}_{ij}) = 0$$

$$\mathcal{H}_k^{(i)}(\mathbf{B}_{ij}) = a_{ijk} > 0, \qquad k \neq i, \qquad k \neq j.$$

For $n = 4$, the configuration of eight points given by

$$\hat{\mathbf{X}}_k = \frac{1}{\sqrt{2}}\left(\cos k\,\frac{\pi}{4}, \sin k\,\frac{\pi}{4}, \cos k\,\frac{\pi}{2}, \sin k\,\frac{\pi}{2}\right) \qquad k = 1, 2, \cdots, 8$$

is a $\{4,8,28\}$. The proof is similar to that just given for the case $n \geqq 5$ with the role of the polynomial $F_{ij}(\lambda)$ being replaced here by the expression

$$F_{ij}^*(\lambda) = \left[1 - \cos\left(\lambda - i\,\frac{\pi}{4}\right)\right] \times \left[1 - \cos\left(\lambda - j\,\frac{\pi}{4}\right)\right].$$

We omit the details.

We close this section by noting that although we cannot show that the biorthogonal code has a largest Q value for codes of $2n$ points, it does have largest nearest neighbor distance, 90° in angular terms. Indeed no collection of more than $n + 1$ vectors in $\mathcal{E}_n$ can have minimum angular distance between points greater than 90°. For consider* Fig. 5. Without

$$\begin{array}{ccccc}
+ & - & - & \cdots & - & - \\
0 & + & - & & - & - \\
0 & 0 & + & & - & - \\
\cdot & \cdot & \cdot & & \cdot & \cdot \\
\cdot & \cdot & \cdot & & \cdot & \cdot \\
\cdot & \cdot & \cdot & & \cdot & \cdot \\
0 & 0 & 0 & & + & -
\end{array}$$

Fig. 5 — Table of component signs.

loss of generality the positive x_1-axis of a rectangular coordinate system can be chosen to lie along the first vector. The first column of the figure shows the sign of the components of this vector. The coordinate axes can be oriented so that $\hat{\mathbf{X}}_2$ lies in the $x_1 - x_2$ plane and the direction of the x_2-axis can be chosen so that the x_2-component of $\hat{\mathbf{X}}_2$ is positive. The second column of Fig. 5 shows the sign of the components of $\hat{\mathbf{X}}_2$. The first component must be negative since if the minimum distance is to be greater than 90° we must have $\hat{\mathbf{X}}_1 \cdot \hat{\mathbf{X}}_2 < 0$. Continuing in this manner we are forced to choose the components of the $\hat{\mathbf{X}}_1, \hat{\mathbf{X}}_2, \cdots, \hat{\mathbf{X}}_{n+1}$ as shown. But now it is impossible to find an $(n + 2)$nd vector having a negative scalar product with these $n + 1$ vectors, for if the nonzero components of $\hat{\mathbf{X}}_{n+2}$ are all negative, it has a positive scalar product with $\hat{\mathbf{X}}_{n+1}$, whereas if the first positive component of $\hat{\mathbf{X}}_{n+2}$ is the jth, $\hat{\mathbf{X}}_j \cdot \hat{\mathbf{X}}_{n+2}$ is positive.

* This elegant proof was suggested by J. H. van Lint.

V. THE INFINITE-BAND GAUSSIAN CHANNEL

When $M = n + 1$ and $f(x)$ is given by (4), the model discussed here describes the transmission of M equally likely signals $s_i(t)$, $i = 1, 2, \cdots, n + 1$, of duration T in white Gaussian noise of spectral power density $N/2$. Here the signals are constrained by

$$\int_0^T s_i^2(t)dt = PT.$$

When these signals are transmitted, the probability of no error using the best possible detection scheme is given by (2), where the $\hat{\mathbf{X}}_i$ must be chosen so that

$$\hat{\mathbf{X}}_i \cdot \hat{\mathbf{X}}_j = \frac{1}{PT} \int_0^T s_i(t)s_j(t)dt,$$

the $\mathcal{R}_i$ are the maximum-likelihood regions (3), and

$$\sigma^2 = \frac{N}{2PT}.$$

See Ref. 3, Sections 4.2 and 4.3 or Ref. 15 for a more detailed description of the correspondence between the geometric model and the physical one.

Our result that the simplex code is optimal means that *in communicating in infinite-band white Gaussian noise by means of* M *equally likely equal-energy signals of duration* T *(no bandwidth restrictions imposed) the error probability is minimized by choosing signals with normalized cross-correlation*

$$\frac{1}{PT} \int_0^T s_i(t)s_j(t)dt = -\frac{1}{M-1}, \qquad i \neq j \tag{25}$$

this being the value of $\hat{\mathbf{X}}_i \cdot \hat{\mathbf{X}}_j$ for the regular simplex.

The error probability with a best set of signals of form (25) is readily determined to be

$$P_e = 1 - \int_{-\infty}^{\infty} dx\, f(x) \Phi^{M-1}\left(x + \frac{1}{\sigma}\sqrt{\frac{M}{M-1}}\right), \tag{26}$$

where $f(x)$ is the Gaussian density (4) and Φ the cumulative

$$\Phi(y) = \int_{-\infty}^{y} f(x)dx.$$

When the transmission rate

$$R = \frac{\log M}{T}$$

is kept constant, along with N and P, (26) becomes for large T (and hence large M)

$$P_e = \exp\,[-E(R)T + o(T)], \tag{27}$$

where

$$E(R) = \begin{cases} \dfrac{C}{2} - R, & R \leqq C/4 \\ (\sqrt{C} - \sqrt{R})^2, & R \geqq C/4 \end{cases} \tag{28}$$

and $C = P/N$ is the capacity of the channel. That the minimal asymptotic error probability for this channel must have the form (27)–(28) was first proved by Wyner.[15]

APPENDIX A

A Lemma

The following lemma will be useful in establishing the main results of Appendices B and C.

Lemma: Let $w_1(x)$ and $w_2(x)$ be integrable functions that satisfy

$$\int_a^b w_1(x)dx = \int_a^b w_2(x)dx. \tag{29}$$

Further, suppose there exists an x', $a \leqq x' \leqq b$, such that

$$\begin{aligned} w_2(x) &\geqq w_1(x), & a \leqq x \leqq x' \\ w_2(x) &\leqq w_1(x), & x' \leqq x \leqq b. \end{aligned} \tag{30}$$

Then, if $m(x)$ is a nonnegative monotone increasing function,

$$\int_a^b m(x)w_1(x)dx \geqq \int_a^b m(x)w_2(x)dx. \tag{31}$$

If $m(x)$ is a nonnegative monotone decreasing function,

$$\int_a^b m(x)w_1(x)dx \leqq \int_a^b m(x)w_2(x)dx. \tag{32}$$

Equality holds in (31) and (32) only if $w_1(x) = w_2(x)$ for almost all x.

Proof: If $m(x)$ is nonnegative and monotone increasing, then

$$
\begin{aligned}
&\int_a^b m(x)[w_1(x) - w_2(x)]dx \\
&\quad = \int_a^{x'} m(x)[w_1(x) - w_2(x)]dx + \int_{x'}^b m(x)[w_1(x) - w_2(x)]dx \\
&\quad \geqq m(x') \int_a^{x'} [w_1(x) - w_2(x)]dx + m(x') \int_{x'}^b [w_1(x) - w_2(x)]dx \\
&\quad = m(x') \left[\int_a^b w_1(x)\,dx - \int_a^b w_2(x)\,dx \right] = 0.
\end{aligned}
$$

If $m(x)$ is nonnegative and monotone decreasing, the steps are the same with the inequalities reversed.

APPENDIX B

Convexity of $h_r(w)$

Let $x_1, x_2, \cdots, x_n$ be the rectangular coordinates of a point in $\mathcal{E}_n$. The surface S of the unit sphere centered at the origin can be given parametrically by

$$
\begin{aligned}
x_1 &= \cos \theta_1 \\
x_2 &= \sin \theta_1 \cos \theta_2 \\
&\vdots \\
x_j &= \sin \theta_1 \sin \theta_2 \cdots \sin \theta_{j-1} \cos \theta_j \\
&\vdots \\
x_{n-1} &= \sin \theta_1 \sin \theta_2 \cdots \sin \theta_{n-2} \cos \theta_{n-1} \\
x_n &= \sin \theta_1 \sin \theta_2 \cdots \sin \theta_{n-2} \sin \theta_{n-1} \\
0 &\leqq \theta_i < \pi, \qquad i = 1, 2, \cdots, n-2 \\
0 &\leqq \theta_{n-1} < 2\pi
\end{aligned}
\tag{33}
$$

and the element of surface content is

$$
ds = \sin^{n-2} \theta_1 \sin^{n-3} \theta_2 \cdots \sin \theta_{n-2}\, d\theta_1\, d\theta_2 \cdots d\theta_{n-1}. \tag{34}
$$

We shall only be concerned with the case $n \geqq 3$.

The spherical cap of angle φ about P, the end point of

$$
\hat{\mathbf{Y}} = (1, 0, 0, \cdots, 0),
$$

is given by $\theta_1 \leqq \varphi$. A hyperplane $\mathcal{H}$ that intersects this spherical cap is

$x_2 = x_1 \tan \alpha$ with $0 \leqq \alpha < \varphi$ and the intersection of $\mathcal{H}$ with the spherical cap is from (33)

$$\cos \theta_2 = \tan \alpha \cot \theta_1 .$$

We then have from the definition (8)

$$h_r(w) = \int_\alpha^\varphi d\theta_1 \int_\mu^\nu d\theta_2 \sin^{n-2} \theta_1 \sin^{n-3} \theta_2 g_r(\sqrt{2 - 2 \cos \theta_1})$$
$$\int_0^\pi d\theta_3 \cdots \int_0^\pi d\theta_{n-2} \int_0^{2\pi} d\theta_{n-1} \sin^{n-4} \theta_3 \cdots \sin \theta_{n-2}$$

where $\nu = \arccos(\tan \alpha \cot \theta_1)$ and $\mu = -\nu$ if $n = 3$ but $\mu = 0$ if $n \geqq 4$. In either event, we can write

$$h_r(w) = k_n \int_\alpha^\varphi d\theta_1 \int_0^\nu d\theta_2 \sin^{n-2} \theta_1 \sin^{n-3} \theta_2 g_r(\sqrt{2 - 2 \cos \theta_1}) \quad (35)$$

while for the content of the piece of the cap cut off by $\mathcal{H}$ we have

$$w = k_n \int_\alpha^\varphi d\theta_1 \int_0^\nu d\theta_2 \sin^{n-2} \theta_1 \sin^{n-3} \theta_2 \quad (36)$$

with $k_n > 0$ and independent of α.

Straightforward differentiation of (35) yields

$$\frac{dh}{d\alpha} = -k_n \sec^2 \alpha \int_\alpha^\varphi d\theta_1 \sin \theta_1 \cos \theta_1$$
$$\cdot [1 - \sec^2 \alpha \cos^2 \theta_1]^{(n-4)/2} g_r(\sqrt{2 - 2 \cos \theta_1}).$$

Now introduce

$$x = \cos^2 \theta_1 , \qquad a = \cos^2 \varphi, \qquad b = \cos^2 \alpha$$

and

$$\hat{g}(x) = g_r(\sqrt{2 - 2\sqrt{x}}).$$

We have $0 \leqq a < b \leqq 1$. Note that $\hat{g}(x) > 0$ is monotone increasing in x. In these terms

$$\frac{dh}{d\alpha} = -\frac{k_n}{2b} \int_a^b dx \left[1 - \frac{x}{b}\right]^{(n-4)/2} \hat{g}(x)$$

and

$$\frac{dw}{d\alpha} = -\frac{k_n}{2b} \int_a^b dx \left[1 - \frac{x}{b}\right]^{(n-4)/2} = -\frac{k_n}{n-2}\left(1 - \frac{a}{b}\right)^{(n-2)/2} .$$

Combining these results we find

$$\frac{dh_r(w)}{dw} = \frac{dh/d\alpha}{(dw/d\alpha)} = \int_a^b l_b(x)\hat{g}(x)dx > 0, \tag{37}$$

where

$$l_b(x) = \frac{\left(1 - \frac{x}{b}\right)^{(n-4)/2}}{\frac{2b}{n-2}\left(1 - \frac{a}{b}\right)^{(n-2)/2}}, \qquad a \leqq x \leqq b. \tag{38}$$

Note that

$$\int_a^b l_b(x)dx = 1. \tag{39}$$

When $n > 4$, the convexity of $h_r(w)$ can be established from (37)–(39) as follows. Consider two different w values, say $w_2 > w_1$ with corresponding parameters $b_2 = \cos^2 \alpha_2$ and $b_1 = \cos^2 \alpha_1$. We have

$$1 \geqq b_2 > b_1 > a.$$

From (38) one readily finds that there is a unique real root x' for which $l_{b_2}(x') = l_{b_1}(x')$, $a < x' < b_1$. For $a \leqq x \leqq x'$ we have $l_{b_1}(x) \geqq l_{b_2}(x)$. If we now define $l_b(x) = 0$ for $x > b$, we can also write $l_{b_2}(x) \geqq l_{b_1}(x)$ for $x \geqq x'$. From (39) we have

$$\int_a^{b_2} l_{b_1}(x)dx = \int_a^{b_2} l_{b_2}(x)dx.$$

The conditions of the lemma of Appendix A hold and we conclude from (37) that

$$w_2 > w_1 \Rightarrow \frac{dh_r(w_2)}{dw} > \frac{dh_r(w_1)}{dw}$$

which is the desired convexity.

When $n = 4$, (38) becomes $l_b(x) = (b - a)^{-1}$ for $a \leqq x \leqq b$. As before, we define $l_b(x) = 0$ for $x > b$. It is readily seen that the lemma again applies with x' chosen as b_1. Convexity is then established in this case as well.

For $n = 3$, (37) and (38) give

$$\frac{dh_r(w)}{dw} = \int_a^b \frac{\hat{g}(x)dx}{2\sqrt{b-a}\sqrt{b-x}}$$
$$= -2\sqrt{b-x}\,\frac{\hat{g}(x)}{2\sqrt{b-a}}\Big|_a^b + \int_a^b \sqrt{\frac{b-x}{b-a}}\,d\hat{g}(x)$$
$$= \hat{g}(a) + \int_a^b \sqrt{\frac{b-x}{b-a}}\,d\hat{g}(x),$$

on integrating by parts. However, since $\hat{g}$ is increasing in x, it follows that the last integral is increasing in b and hence also in w. The convexity proof is thus completed.

APPENDIX C

Proof of Equation (11)

We shall be concerned here with two different regions, V and W, cut off from the spherical cap of angle φ about the point P which we take as the terminus of the unit vector $\hat{\mathbf{Y}}$ in $\mathcal{E}_n$ (see Fig. 6). The region V is the intersection of the spherical cap with a convex cone $\mathcal{V}$ having the origin as a vertex. It is assumed that $\mathcal{V}$ does not contain P. We denote by Q a point of V closest to P. The second region, W, is cut off from the cap by a single hyperplane $\mathcal{L}$ through the origin but not through P. $\mathcal{L}$ is chosen so that W and V have the same content, w and v, respectively, and for purposes of our proof we restrict the normal to $\mathcal{L}$ to lie in the 2-plane through the origin, P and Q. We wish to show that

$$I_V \equiv \int_V g_r(|\,\hat{\mathbf{Y}} - \hat{\mathbf{X}}\,|)ds \geqq I_W \equiv \int_W g_r(|\,\hat{\mathbf{Y}} - \hat{\mathbf{X}}\,|)ds \qquad (40)$$

with equality holding only if V is cut off from the cap by a single hyperplane. Here, as in (11), g_r is nonnegative and monotone decreasing and $\hat{\mathbf{X}}$ is a generic unit vector in $\mathcal{E}_n$. In the applications made of (40) in the main text, $\mathcal{V}$ is specialized to a type of flat-sided cone.

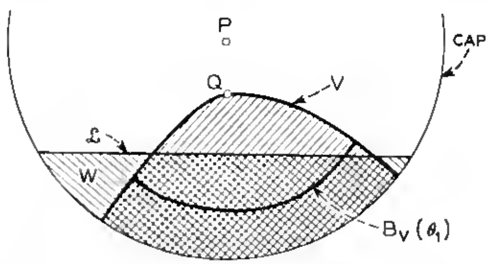

Fig. 6 — Regions involved in proof of (11).

Let us again adopt the spherical coordinates (33) with the pole P of the cap located on the x_1-axis so that $\hat{\mathbf{Y}} = (1,0,0, \cdots, 0)$. We suppose the axes are oriented so that Q lies in the x_1-x_2 plane. Then from (33) and (34)

$$I_V = \int_0^{\varphi} d\theta_1 g_r(\sqrt{2 - 2\cos\theta_1})\, v(\theta_1) \tag{41}$$

where

$$v(\theta_1) = \sin^{n-2}\theta_1 \underset{B_V(\theta_1)}{\int d\theta_2 \cdots \int} d\theta_{n-1} \sin^{n-3}\theta_2 \cdots \sin\theta_{n-1} \tag{42}$$

is the $(n-2)$-dimensional content of the intersection $B_V(\theta_1)$ of V with the hyperplane $x_1 = \cos\theta_1$. Similarly,

$$I_W = \int_0^{\varphi} d\theta_1 g_r(\sqrt{2 - 2\cos\theta_1})\, w(\theta_1) \tag{43}$$

where $w(\theta_1)$ is the $(n-2)$-dimensional content of the intersection $B_W(\theta_1)$ of W with the hyperplane $x_1 = \cos\theta_1$. By hypothesis we have

$$v = \int_0^{\varphi} d\theta_1 v(\theta_1) = w = \int_0^{\varphi} d\theta_1 w(\theta_1). \tag{44}$$

Since $g_r(\sqrt{2 - 2\cos\theta_1})$ is a nonnegative monotone decreasing function of θ_1, all the hypotheses of the lemma of Appendix A will hold if we can show the existence of a φ' such that

$$\begin{aligned} v(\theta_1) &\geqq w(\theta_1), & 0 \leqq \theta_1 &\leqq \varphi' \\ w(\theta_1) &\geqq v(\theta_1), & \varphi' \leqq \theta_1 &\leqq \varphi. \end{aligned} \tag{45}$$

The conclusion (32) of the lemma then is (40).

Our goal now, therefore, is to show that $v(\theta)$ and $w(\theta)$ cross only once as indicated in (45). Let $\alpha = \angle POQ$. If Q^* is the nearest point in W to P and $\beta = \angle POQ^*$, then $\beta > \alpha$. For $0 \leqq \theta \leqq \alpha$, both $v(\theta)$ and $w(\theta)$ are zero. For $\alpha < \theta \leqq \beta$, $v(\theta) > w(\theta) = 0$. From (44) it then follows that there is a first point in $(0,\varphi)$ where $w(\theta)$ crosses up through $v(\theta)$, that is, where $v(\theta) = w(\theta)$ and $w'(\theta) > v'(\theta)$ where the prime denotes differentiation with respect to θ. If there were a second crossing, at that point we would have $w' < v'$. We prove that there is only one crossing by demonstrating that

$$v(\theta_1) = w(\theta_1) \Rightarrow \frac{dw(\theta_1)}{d\theta_1} \geqq \frac{dv(\theta_1)}{d\theta_1} \tag{46}$$

for $0 \leqq \theta_1 \leqq \varphi$.

Let ω be such that $v(\omega) = w(\omega)$, $0 \leqq \omega \leqq \varphi$. Consider now the spherical pyramid Γ_V having Q as vertex and as base the set $B_V(\omega)$ defined below (42). Γ_V is the set of all points $\hat{\mathbf{X}}$ of the form

$$\begin{gathered} \hat{\mathbf{X}} = \xi \hat{\mathbf{X}}_Q + \eta \hat{\mathbf{X}}_B, \\ \xi \geqq 0, \qquad \eta \geqq 0, \qquad \hat{\mathbf{X}}_Q = \overrightarrow{OQ}, \qquad \hat{\mathbf{X}}_B \in B_V(\omega) \end{gathered} \tag{47}$$

where, in order for (47) to be a unit vector, we have the additional restriction

$$|\hat{\mathbf{X}}|^2 = 1 = \xi^2 + \eta^2 + 2\xi\eta \hat{\mathbf{X}}_Q \cdot \hat{\mathbf{X}}_B. \tag{48}$$

Note that since $\mathcal{V}$ is convex and since Q and $B_V(\omega)$ are contained in $\mathcal{V}$, it follows from (47) that Γ_V is contained in $\mathcal{V}$ and S, hence Γ_V is contained also in V.

Now let $\bar{v}(\theta_1)$ denote the $(n - 2)$-dimensional content of the intersection of Γ_V with the hyperplane $x_1 = \cos\theta_1$, where $\alpha \leqq \theta_1 \leqq \omega$. We have

$$\begin{aligned} \bar{v}(\omega) &= v(\omega) \\ \bar{v}(\omega - \delta) &\leqq v(\omega - \delta) \end{aligned} \tag{49}$$

where this last follows from the fact that Γ is contained in V. One has then

$$\frac{v(\omega) - v(\omega - \delta)}{\delta} \leqq \frac{\bar{v}(\omega) - \bar{v}(\omega - \delta)}{\delta}$$

so that

$$\left.\frac{dv(\theta_1)}{d\theta_1}\right|_{\theta_1=\omega} \leqq \left.\frac{d\bar{v}(\theta_1)}{d\theta_1}\right|_{\theta_1=\omega}. \tag{50}$$

Consider next the spherical pyramid $\Gamma_W(Q)$ having Q as vertex and as base the set $B_W(\omega)$ defined below (43). We denote by $\bar{w}(\theta_1)$ the $(n - 2)$-dimensional content of the intersection of Γ_W with the hyperplane $x_1 = \cos\theta_1$. As before, let Q^* be the nearest point in W to P. We denote by $\bar{w}^*(\theta_1)$ the $(n - 2)$-dimensional content of the intersection of the spherical pyramid $\Gamma_W(Q^*)$ with the hyperplane $x_1 = \cos\theta_1$. Since Q^* is contained in $\Gamma_W(Q)$, $\Gamma_W(Q^*)$ is also contained in $\Gamma_W(Q)$ and we have

$$\begin{aligned} \bar{w}(\omega) &= \bar{w}^*(\omega) \\ \bar{w}(\omega - \delta) &\geqq \bar{w}^*(\omega - \delta). \end{aligned} \tag{51}$$

From this it follows that

$$\left.\frac{d\bar{w}(\theta_1)}{d\theta_1}\right|_{\theta_1=\omega} \leqq \left.\frac{d\bar{w}^*(\theta_1)}{d\theta_1}\right|_{\theta_1=\omega}.$$

However, since Q^* lies in the hyperplane,

$$\bar{w}^*(\omega) = w(\omega) \tag{52}$$

for all ω, hence

$$\left.\frac{d\bar{w}(\theta_1)}{d\theta_1}\right|_{\theta_1=\omega} \leqq \left.\frac{dw(\theta_1)}{d\theta_1}\right|_{\theta_1=\omega}. \tag{53}$$

In the remaining paragraphs of this appendix we shall show that

$$\bar{v}(\omega) = \bar{w}(\omega) \Rightarrow \left.\frac{d\bar{v}(\theta_1)}{d\theta_1}\right|_{\theta_1=\omega} \leqq \left.\frac{d\bar{w}(\theta_1)}{d\theta_1}\right|_{\theta_1=\omega} \tag{54}$$

which will establish (46) and complete our proof, for the hypothesis of (46) follows from that of (54) by (49), (51), and (52) and the conclusion of (46) follows from the conclusion of (54) by (50) and (53).

Let the spherical coordinates of a point $\hat{\mathbf{X}}$ in Γ_V be denoted by the angles $(\varphi_1, \cdots, \varphi_{n-1})$ (see Fig. 7). We employ the angles $(\theta_1, \cdots, \theta_{n-1})$ to describe a point $\hat{\mathbf{X}}_B$ in $B_V(\theta_1)$. The content $\bar{v}(\mu)$ of the intersection $J(\mu)$ of Γ_V with the byperplane $x_1 = \cos\mu$, $\alpha \leqq \mu \leqq \omega$ is given by

$$\bar{v}(\mu) = \sin^{n-2}\mu \int d\varphi_2 \sin^{n-3}\varphi_2 \int d\varphi_3 \sin^{n-4}\varphi_3 \cdots \int_{J(\mu)} d\varphi_{n-1}. \tag{55}$$

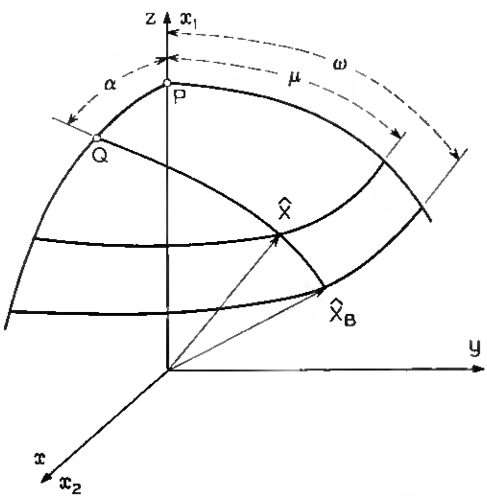

Fig. 7 — The mapping from $\hat{\mathbf{X}}$ to $\hat{\mathbf{X}}_B$.

The relationships (47)-(48), however, serve to define a one-one transformation between the coordinates $(\mu, \varphi_2, \cdots, \varphi_{n-1})$ of a point in $J(\mu)$ and the coordinates $(\omega, \theta_2, \cdots, \theta_{n-1})$ of a point in $B_V(\omega)$, so that $\bar{v}(\mu)$ can be expressed as an integration over $B_V(\omega)$ as well. Taking components of (47), we find successively

$$
\begin{aligned}
\cos \mu &= \xi \cos \alpha + \eta \cos \omega \\
\sin \mu \cos \varphi_2 &= \xi \sin \alpha + \eta \sin \omega \cos \theta_2 \\
\sin \mu \sin \varphi_2 \cos \varphi_3 &= \eta \sin \omega \sin \theta_2 \cos \theta_3 \\
&\vdots \\
\sin \mu \sin \varphi_2 \cdots \sin \varphi_{j-1} \cos \varphi_j &= \eta \sin \omega \sin \varphi_2 \cdots \sin \varphi_{j-1} \cos \varphi_j \qquad (56) \\
&\vdots \\
\sin \mu \sin \varphi_2 \cdots \sin \varphi_{n-2} \sin \varphi_{n-1} & \\
&= \eta \sin \omega \sin \varphi_2 \cdots \sin \varphi_{n-2} \sin \varphi_{n-1} \\
j &= 3, 4, \cdots, n-1.
\end{aligned}
$$

Dividing the nth equation by the $(n-1)$st yields $\varphi_{n-1} = \theta_{n-1}$. Dividing the $(n-1)$st equation by the $(n-2)$nd then yields $\varphi_{n-2} = \theta_{n-2}$. Proceeding in this manner, one finds $\varphi_j = \theta_j$, $j = 3, 4, \cdots, n-1$. The first two equations of (56) can be solved for ξ and η. By substituting these expressions into (48) which now reads

$$\xi^2 + \eta^2 + 2\xi\eta\,(\cos \alpha \cos \omega + \sin \alpha \sin \omega \cos \theta_2) = 1, \tag{57}$$

we obtain a single relationship connecting φ_2 and θ_2 which we suppose solved in the form

$$\varphi_2 = \varphi_2(\theta_2, \mu). \tag{58}$$

Equation (55) now becomes in the new variables

$$
\begin{aligned}
\bar{v}(\mu) &= \sin^{n-2} \mu \underset{B_V(\omega)}{\int d\theta_2 \frac{d\varphi_2}{d\theta_2} \sin^{n-3} \varphi_2 \int d\theta_3 \sin^{n-2} \theta_3 \cdots \int d\theta_{n-1}} \\
&= \sin^{n-2} \omega \int d\theta_2 G(\theta_2, \mu) h(\theta_2)
\end{aligned} \tag{59}
$$

with

$$G(\theta_2, \mu) = \left[\frac{\sin \mu}{\sin \omega}\right]^{n-2} \frac{d\varphi_2}{d\theta_2} \left[\frac{\sin \varphi_2}{\sin \theta_2}\right]^{n-3} \tag{60}$$

and

$$h(\theta_2) = \sin^{n-3}\theta_2 \int\limits_{B_V(\omega)} d\theta_3 \sin^{n-4}\theta_3 \cdots \int d\theta_{n-1}, \tag{61}$$

where if $n = 3$ this latter expression is to he interpreted as unity. It is convenient now to define $h(\theta_2)$ to he zero if θ_2 is not the second angle coordinate of a point in $B_V(\omega)$. In this notation, then, we have for $n \geqq 4$

$$\left.\frac{d\bar{v}(\mu)}{d\mu}\right|_{\mu=\omega} = \sin^{n-2}\omega \int_0^\pi d\theta_2 \left.\frac{\partial G(\theta_2,\mu)}{\partial\mu}\right|_{\mu=\omega} h(\theta_2) \tag{62}$$

$$\bar{v}(\omega) = \sin^{n-2}\omega \int_0^\pi d\theta_2 h(\theta_2). \tag{63}$$

If $n = 3$, the lower limits of integration here should he replaced hy $-\pi$. It will he shown later that $\partial G/\partial\mu \,|_{\mu=\omega}$ is a nonnegative monotone decreasing function of θ_2.

We next seek to determine the nature of the set $B_V(\omega)$ of given content $\bar{v}(\omega)$ that will maximize (62). We note first from (61) that for $n \geqq 4$

$$h(\theta_2) \leqq \sigma(\theta_2) \tag{64}$$

where $\sigma(\theta_2)$ is the surface content of a sphere of radius $\sin\theta_2$ in $\mathcal{E}_{n-2}$ since σ is given by the integrals of (61) with the integration variables running through their maximum allowahle range. Now let $B^*(\omega)$ he the set of points defined hy $\theta_1 = \omega$, $0 \leqq \theta_2 \leqq \theta_2'$ where θ_2' is given hy

$$\bar{v}(\omega) = \sin^{n-2}\omega \int_0^{\theta_2'} d\theta_2 \sigma(\theta_2).$$

For $B^*(\omega)$ we have

$$h^*(\theta_2) = \begin{cases} \sigma(\theta_2), & 0 \leqq \theta_2 \leqq \theta_2' \\ 0, & \theta_2' < \theta_2 \end{cases} \tag{65}$$

so that

$$\bar{v}(\omega) = \sin^{n-2}\omega \int_0^\pi d\theta_2 h^*(\theta_2). \tag{66}$$

We also have

$$\begin{aligned} h^*(\theta_2) &\geqq h(\theta_2), & 0 \leqq \theta_2 \leqq \theta_2' \\ h^*(\theta_2) &\leqq h(\theta_2), & \theta_2' \leqq \theta_2 \leqq \pi \end{aligned} \tag{67}$$

from (64) and (65).

The hypotheses of the lemma of Appendix A are thus met from (63), (66), (67), and the monotonicity of $\partial G/\partial\mu$. We conclude that among sets of equal content, $d\bar{v}/d\mu\,|_{\mu=\omega}$ is a maximum for the set $B^*(\omega)$. The set $B_W(\omega)$, however, coincides with $B^*(\omega)$. Equation (54) then follows for $n \geqq 4$. The modification necessary to treat the case $n = 3$ is trivial.

There remains only the demonstration that $\partial G/\partial\mu\,|_{\mu=\omega}$ is nonnegative monotone decreasing in θ_2. Equation (57) and the first two equations of (56) are identical with the equations that would hold for the three 3-vectors $\overrightarrow{OQ}$, $\hat{\mathbf{X}}$ and $\hat{\mathbf{X}}_B$ of Fig. 7 constrained to satisfy (47). The relationship (58) between φ_2 and θ_2 can most easily be written down by consulting this figure. The condition that the three points be coplanar with the origin is

$$\begin{vmatrix} x & y & z \\ x_Q & y_Q & z_Q \\ x_B & y_B & z_B \end{vmatrix} = 0 \tag{68}$$

where

$$\begin{aligned} x &= \sin\mu\cos\varphi_2 & y &= \sin\mu\sin\varphi_2 & z &= \cos\mu \\ x_Q &= \sin\alpha & y_Q &= 0 & z_Q &= \cos\alpha \\ x_B &= \sin\omega\cos\theta_2 & y_B &= \sin\omega\sin\theta_2 & z_B &= \cos\omega \end{aligned} \tag{69}$$

which serves to determine (58). Routine implicit differentiation of (68) and (69) and evaluation at $\mu = \omega$, $\varphi_2 = \theta_2$ yields

$$\left.\frac{d\varphi_2}{d\theta_2}\right|_{\mu=\omega} = 1 \tag{70}$$

$$\left.\frac{\partial\varphi_2}{\partial\mu}\right|_{\mu=\omega} = \frac{\sin\alpha\sin\theta_2}{\sin\omega(\cos\alpha\sin\omega - \sin\alpha\cos\omega\cos\theta_2)} \tag{71}$$

$$\left.\frac{\partial}{\partial\mu}\frac{d\varphi_2}{d\theta_2}\right|_{\mu=\omega} = \frac{\sin\alpha[\cos\alpha\sin\omega\cos\theta_2 - \sin\alpha\cos\omega]}{\sin\omega[\cos\alpha\sin\omega - \sin\alpha\cos\omega\cos\theta_2]^2}. \tag{72}$$

The denominators of (71) and (72) are positive since $\omega > \alpha$ implies $\tan\omega\cot\alpha > 1 \geqq \cos\theta_2$ which is the same as

$$\cos\alpha\sin\omega > \sin\alpha\cos\omega\cos\theta_2\,.$$

The numerator of (72) is nonnegative for points $\hat{\mathbf{X}}_B$ of interest to us since we are concerned only with points in the portion of the cap cut off by the hyperplane that passes through Q and througb tbe origin O and

has its normal lying in the plane POQ; i.e., points for which $x_2 \geqq x_1 \tan \alpha$. For points in this region on the sphere and in the hyperplane $x_1 = \cos \omega$ this inequality is

$$\sin \omega \cos \theta_2 \geqq \cos \omega \tan \alpha$$

or

$$\cos \alpha \sin \omega \cos \theta_2 - \sin \alpha \cos \omega \geqq 0.$$

Now from (60) and (70)

$$\left.\frac{\partial G}{\partial \mu}\right|_{\mu=\omega} = (n-2)\frac{\cos \omega}{\sin \omega} + \left.\frac{\partial}{\partial \mu}\frac{d\varphi_2}{d\theta_2}\right|_{\mu=\omega} + (n-3)\left.\frac{\cos \theta_2}{\sin \theta_2}\frac{\partial \varphi_2}{\partial \mu}\right|_{\mu=\omega}.$$

Using (71) and (72), it is readily seen that this expression is nonnegative and monotone decreasing in θ_2.

REFERENCES

1. Shannon, C. E., Communication in the Presence of Noise, Proc. IRE, *37*, January, 1949, pp. 10–21.
2. Kotel'nikov, V. A., thesis, Molotov Energy Institute, Moscow, 1947, translated by R. A. Silverman as *The Theory of Optimum Noise Immunity*, McGraw-Hill Book Co., New York, 1959.
3. Wozencraft, J. M. and Jacobs, I. M., *Principles of Communication Engineering*, John Wiley & Sons, New York, 1965.
4. Shannon, C. E., Probability of Error for Optimal Codes in a Gaussian Channel. B.S.T.J., *38*, May, 1959, pp. 611–656.
5. Gallager, R. G., A Simple Derivation of the Coding Theorem and Some Applications, IEEE Trans., *IT-11*, January, 1965, pp. 3–18.
6. Slepian, D., Bounds on Communication, B.S.T.J., *42*, May, 1963, pp. 681–707.
7. MacColl, L. A., Signalling in the Presence of Thermal Noise, I, II, and III, Bell Laboratories internal memoranda issued May 27, June 30, and September 13, 1948.
8. Coxeter, H. S. M., *Regular Polytopes*, MacMillan Co., New York, 1963.
9. Gilbert, E. N., A Comparison of Signalling Alphabets, B.S.T.J., *31*, May, 1952, pp. 504–522.
10. Balakrishnan, A. V., A Contribution to the Sphere-Packing Problem of Communication Theory, J. Math. Anal. Appl., *3*, No. 3, December, 1961, pp. 485–506.
11. Balakrishnan, A. V., Signal Selection Theory for Space Communication Channels, Chapter 1 in *Advances in Communication Systems*, A. V. Balakrishnan, editor, Academic Press, New York, 1965.
12. Weber, C. L., On Optimum Signal Selection for M-ary Alphabets with Two Degrees of Freedom, IEEE Trans., *IT-11*, No. 2, April, 1965, pp. 299–300; New Solutions to the Signal Design Problem for Coherent Channels, IEEE Trans., *IT-12*, No. 2, April, 1966, pp. 161–167.
13. Fejes Tóth, L., *Lagerungen in der Ebene auf der Kugel und im Raum*, Springer-Verlag, Berlin, 1953, pp. 137–138.
14. Gale, D., Neighborly and Cyclic Polytopes, Proc. Symposia Pure Math., *VII*, pp. 225–232, Am. Math. Soc., Providence, Rhode Island, 1963.
15. Wyner, A., On the Probability of Error for Communication in White Caussian Noise, to appear IEEE Trans., *IT*-.